Volume Title: Annals of Economic and Social Measurement, Volume 1, number 3

Volume Author/Editor: NBER

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm72-3

Publication Date: July 1972

Chapter Title: Research Methodology Notes: The Inadequacy of Cross-Section Age-Earnings Profiles When Ability Is Not Held Constant

Chapter Author: Paul Taubman, Terence Wales

Chapter URL: http://www.nber.org/chapters/c9436

Chapter pages in book: (p. 363 - 368)

Annals of Economic and Social Measurement, 1/3, 1972

RESEARCH METHODOLOGY NOTES

THE INADEQUACY OF CROSS-SECTION AGE-EARNINGS PROFILES WHEN ABILITY IS NOT HELD CONSTANT

BY PAUL TAUBMAN AND TERENCE WALES

Cross-section samples such as the Census are often used to construct age-earnings profiles at different education levels. In constructing such profiles it is assumed, among other things, that within an education group, each age group or cohort is identical with respect to all characteristics that determine income. To help achieve this homogeneity of cohorts, profiles are sometimes estimated from regressions which in addition to education and age, variables such as location, marital status, health, etc. have been included. However data are never available on all individual characteristics that determine earnings. In this note we examine the effects of omitting mental ability on both the shape of the age-earnings profile, and on the differences between profiles.

One would suspect a priori that the omission of mental ability measures would be particularly serious in estimating the extra earnings due to education since such abilities are probably an important determinant of earnings, and also highly correlated with education.[1] Hence as shown formally below estimated returns to education from cross-section data such as those provided in the Census, for which no measure of ability is available, will be biased upwards. It is a purpose of this note to suggest that not only will such returns be biased upward, but that the extent of the bias will depend on the particular cohort studied. Hence a whole set of bias corrections is required (and not just a single one) in order to obtain unbiased estimates of the returns to education.[2] Moreover because of the particular cohorts involved, an average bias correction in recent Census data will lead to an underestimate in the rate of return to higher education. Finally since the bias has varied over time, unqualified comparisons of rates of return from the 1939, 1949, 1959, etc. Census are not valid. To complicate matters even more, education and ability coefficients do not change proportionately over time; hence, the slope of the profile and the bias correction also depend on the age at which a particular cohort is studied.

Bias from Omitting Ability

Suppose for simplicity that the true relationship between earnings (Y), education (S) and mental ability (A) is given by:

$$Y = b_0 + b_1 A + b_2 S + u \tag{1}$$

where u is a random disturbance independent of the explanatory variables. Further consider the linear relation between ability and education given by:

$$A = c_0 + c_1 S + w \tag{2}$$

[1] See Taubman–Wales [3] for some estimates of the size of the bias.

[2] These corrections are in addition to those arising from the possibility that the bias varies with the individual's age and/or education level attained. See below.

where w is again a random error. Then least-squares estimation of equation 1 with ability omitted provides an estimate of b_2 with the property that its expectation equals $b_2 + b_1c_1$. Hence as long as ability positively influences income ($b_1 > 0$) and education and ability are positively related ($c_1 > 0$), the estimate of b_2 will be biased upward in the absence of data on ability.

The coefficient c_1 in the relationship between ability and education is an integral part of the bias. Suppose that returns to education are now estimated *with ability omitted* using a cross-section sample involving various cohorts of individuals (e.g. those aged 25–35, 35–45, 45–55, etc.). If the relationship given in equation 2 has changed over time then the bias on the education coefficient (b_1c_1) will differ between cohorts. But, since individuals in different cohorts were educated at various times in the past, perhaps from 0 up to 50 years ago, it is likely that such differences exist. This is especially true since equation 2, a descriptive reduced form relationship between ability and education, embodies changes on both the demand and supply side for education. The enormous increases in the demand for education during the twentieth century together with increases in the supply of facilities, the changing nature of higher education itself, and increased financial assistance to students are some of the factors that may have affected this relationship. It is not necessary to speculate on this issue, however, since from information in [4] we can estimate c_1 for various cohorts.

In [4] we analyzed information from various samples spanning the twentieth century in an attempt to shed some light on the behavior of the ability-education relation over time.[3] Necessarily we were restricted to simple measures of ability and education—the percent of high school graduates who enter college at various IQ levels. In order to obtain comparable results from different samples, we converted the IQ ability measures in each sample into percentile terms, with the "norm" being the population of high school graduates. In order to compare and combine samples from different time periods we assume that the average ability level of high school graduates remained approximately constant over the time period. Support for this hypothesis is contained in [1] and [2].

In [4] we were able in some instances to obtain separate estimates of equation 2 for males and for females. But because such separation was not available for cohorts finishing high school in the 1920's, we will use equations in which data for both genders are combined. In [4] we present estimates of equation 2 for many samples. To save space we will not repeat the equations but merely indicate in Table 1 the estimate of c_1 in each sample.[4]

To test whether c_1 differed significantly between samples, we estimated for each pair of samples an equation of the form:

(3)
$$A = c_0 + d_0D + (c_1 + d_1D)E$$

where A is the IQ percentile, E is the percent of high school graduates at each level who entered college, and D is a dummy variable taking on values of 1 for one

[3] The samples used and a more detailed explanation of some aspects of the procedure discussed below are described in detail in [4]. The samples are for the years 1925, 1929, 1934, 1938, 1946, 1950, 1957, 1960, and 1961.

[4] Of course in studying the returns to education one generally uses various education categories, e.g., high school, B.A., M.A., etc. Unfortunately no data are available over time on these education breakdowns by ability level and we are forced to use this one measure of educational achievement as a proxy for the behavior of others when we apply the results to earnings equations.

of the samples and zero for the other. The c's and d's are coefficients to be estimated. We are interested in determining whether the slope coefficient in equation 3 differs significantly between samples, that is whether d_1 is significantly different from zero.

TABLE 1

COMPARISON OF SLOPE COEFFICIENTS IN REGRESSIONS OF ABILITY ON EDUCATION
(t values for differences in slope coefficients)

	1925 (2.72)	1929 (2.24)	1934 (1.87)	1938 (2.36)	1946 (1.42)	1950 (1.61)	1957 (1.55)	1960 (1.17)	1961 (1.29)
1925	—	—	—	—	—	—	—	—	—
1929	0.8	—	—	—	—	—	—	—	—
1934	1.5	0.5	—	—	—	—	—	—	—
1938	0.3	0.3	1.2	—	—	—	—	—	—
1946	2.8	1.4	1.2	3.0	—	—	—	—	—
1950	2.1	0.9	0.8	2.1	0.8	—	—	—	—
1957	1.7	0.7	0.7	1.7	0.4	0.5	—	—	—
1960	4.2	2.3	2.9	5.3	1.4	4.0	2.8	—	—
1961	2.6	1.3	1.7	2.9	0.1	3.0	2.0	1.3	—

Note: The regression equations are of the form: $A = c_0 + d_0D + (c_1 + d_1D)E$, where A is an IQ percentile and E is the fraction of high school graduates entering college at that IQ level. Table entries are t values for d_1 for all pairs of samples. Coefficients in parentheses at the top of the table are the slope coefficients (expressible either as c_1 or $c_1 + d_1$) for each sample.

Table 1 contains the estimated t values for the coefficient d_1 from various pairs of samples, as well as the coefficients for the various years in which the samples were drawn. The slope coefficients, which show a marked downward trend over the period, can be divided into three to four distinct groups. Only the 1930's do not differ significantly from other periods, and it is not clear whether they should be combined with the 1920's, the immediate postwar period, or be considered as a separate group. We have combined the samples within broader periods and have reestimated the equations to obtain the following results. There has been a continual decrease in the slope coefficient between the 1920's and 1960's with the latter value only about one-half of the former.[5]

TABLE 2

SLOPE COEFFICIENTS IN REGRESSIONS OF ABILITY ON EDUCATION

Time Period	Slope Coefficient
1920's	2.53
1920's and 1930's	2.03
1930's	1.87
1930's and 1950's	1.57
1950's	1.56
1960's	1.27

[5] However, data available for cohorts of males whose education ended prior to World War I indicate a slope coefficient equal to that of males in the 1960's.

The above findings lead to several interesting conclusions. First and not surprisingly, there is a correlation between ability and education which as shown in [4] is significant. Thus omission of ability will bias education coefficients provided ability influences earnings. Second the (portion of the) bias correction due to c_1 will vary for high school graduates in different cohorts. The inappropriateness of using a single correction fo· the bias in education coefficients is reinforced by the finding in [3] that the bias from omitting ability is less (in percentage terms) when the same people are 47 years old than when they are 33.[6]

Another important conclusion is that the bias on the rate of return to education calculated from Census data for 1939, 1949, etc. is not the same. Hence the constancy of the rate of return using these Census data does not indicate that the unbiased rate of return has remained constant. Finally, the use of an average correction for the omission of ability leads to an underestimate of the rate of return to education when the bias is smaller for younger cohorts, because discounting reduces more the weight given to earnings differences of the elderly.

Age-Earning Profiles for a Given Education Level

Several issues in human capital theory hinge on how mean earnings and the distribution of such earnings vary with age at different education levels. Census data are often used to estimate average earnings profiles and the distribution about the average. Such calculations must assume that within an education level people at all ages have the same average ability. The previous discussion focused only on changes in the marginal relationship between education and mental ability. However in [4] we also calculated the average ability of those high school graduates who did and did not enter college. The results, presented in Figure 1, indicate quite clearly that the average ability level of college entrants has risen and non-college entrants has fallen since the 1920's.[7] Since younger cohorts of the more educated are more able, Census cross-section age-income profiles will be too flat at high education levels and too steep at low education levels provided ability affects income.

There is still another difficulty in comparing the steepness (or growth rate) of the profiles at various education levels. From equation 1 it is clear that given education, average earnings depend on b_1 times the average ability level. Both b_1 and b_2 vary with age, but at least for the portion of the population studied in [3] the effect of education increases more from age 33 to 47 than the effect of ability. Since the average ability level is correlated with education, the growth in average earnings by education level understates the difference in earnings due solely to education. For example in [3] we calculated the growth rate in earnings for individuals between the ages of 33 and 47 using first average earnings and then earnings of a person with a given set of socio-economic characteristics. In Table 3 we present the relative growth in earnings, i.e. the growth at any education level minus the

[6] As the people age, b_1 and b_2 in equation 1 increase while c_1 is unchanged. Since b_1 grows less than b_2 the bias (in percentage terms) decreases.

[7] The numbers in Figure 1 have been adjusted to equate the average percent of students who entered college in a state with the average for the nation in that year. See Appendix C of [4] for details. The same pattern appears though with less precision if no adjustment is made.

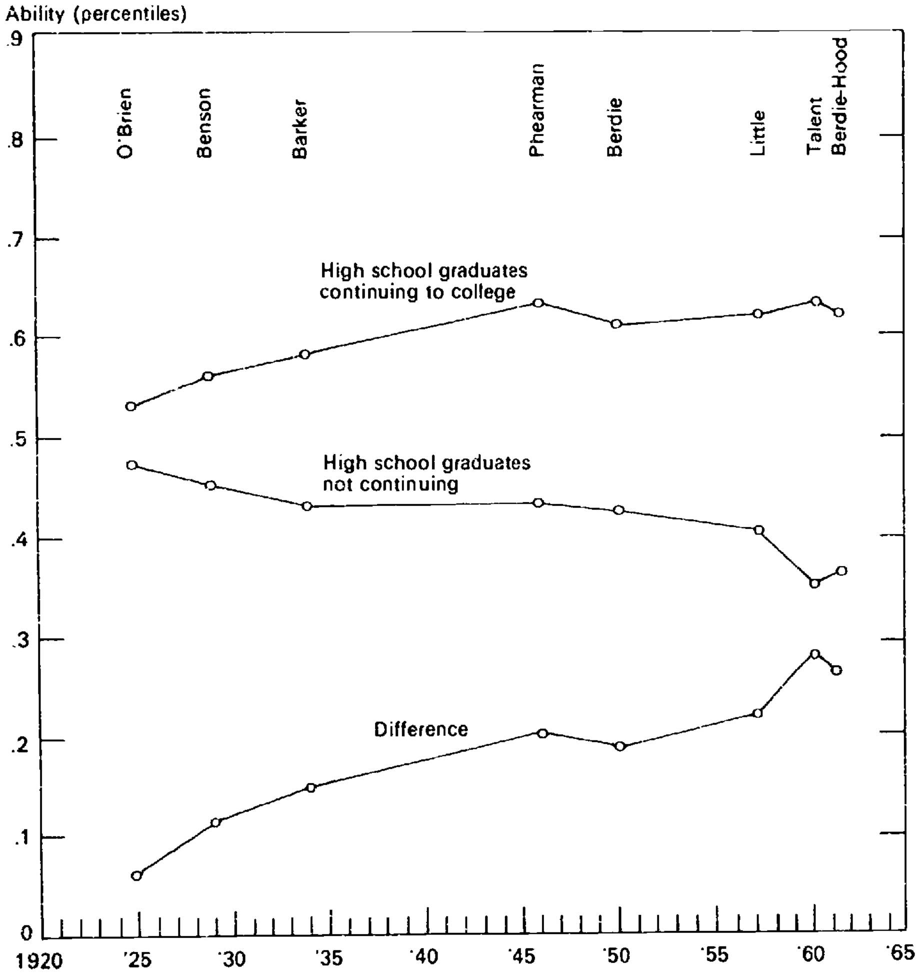

Figure 1 Average Ability Levels Over Time, Adjusted

TABLE 3

PERCENTAGE GROWTH IN EARNINGS BY EDUCATION LEVEL RELATIVE TO GROWTH FOR HIGH SCHOOL GRADUATES, 1955–69

Education Level	Mean Earnings	Mean Earnings after Standardizing for Mental and Other Sociodemographic Characteristics
Some college	55	60
Undergraduate degree	158	104
Some graduate	72	72
Masters	220	166
Ph.D.	1,250	550
M.D.	47	51
L.L.B.	342	292

Source: [3], Chapter 6, Table 2B.

growth at the high school level calculated on the basis of both of these assumptions. For the 100 Ph.D.'s the ratio of the two measures is 2.3. Even for the 1,100 people with a bachelors degree, the ratio is 1.5.

It seems clear, therefore, that Census data which do not include ability measures are inappropriate for studying age earnings profiles.

CONCLUSION

In this paper we have examined the consequences of using Census data in which there is no information on mental ability to answer certain questions posed by the theory of human capital. Drawing on previous work in [4] we have shown that the average and marginal relationships between higher education and mental ability have varied by cohorts. Further, information in [3] indicates that ability is an important determinant of earnings, and that for the same group of people the effects of education and ability do not increase at the same rate with age. Thus different corrections should be applied to differences in earnings between age levels both within education groups and between education groups. Alternatively, age-earnings profiles should be calculated for people at given ability as well as education levels.

In short, it is a poor practice to use information on earnings in different cohorts from a cross-section sample to study questions in the theory of human capital if ability measures are not available.

Wharton School of Finance and Commerce,
University of Pennsylvania, and NBER,
and
University of British Columbia and NBER

REFERENCES

[1] Berdie, et al. *Who Goes to College? Comparison of Minnesota College Freshmen, 1930–60*, University of Minnesota, Minneapolis, 1962.

[2] Finch, F. H., "Enrollment Increases and Changes in the Mental Level of the High School Population," *Applied Psychology Monograph*, 1946, No. 10.

[3] Taubman, P. and Wales, T. *Education as an Investment and Screening Device*, University of Pennsylvania and University of British Columbia (manuscript), 1971.

[4] Taubman, P. and Wales, T. "Mental Ability and Higher Educational Attainment in the Twentieth Century." NBER *Occasional Paper* 118 and Carnegie Commission on Higher Education Report, 1972.